Comparison of energy expenditure estimates from 4 physical activity questionnaires with doubly labeled water estimates in postmenopausal women[1–3]

Somdat Mahabir, David J Baer, Carol Giffen, Beverly A Clevidence, William S Campbell, Philip R Taylor, and Terry J Hartman

ABSTRACT

Background: Physical activity energy expenditure (EE) is an important determinant of health, and epidemiologists have used various methods, such as physical activity and energy intake recalls and records, to estimate energy cost. However, most epidemiologic studies have not validated these methods against the doubly labeled water (DLW) technique for measuring EE.

Objective: The aim was to compare EE estimated by 4 physical activity questionnaires with that obtained with the DLW technique in free-living postmenopausal women.

Design: We measured EE in kcal/d using the DLW method, the Harvard Alumni questionnaire, the Five City Project questionnaire, the Cross-Cultural Activity Participation Study (CAPS) Four Week Activity Recall, and the CAPS Typical Week Activity Survey in 65 healthy postmenopausal women.

Results: Compared with DLW, the Harvard Alumni questionnaire, the Five City Project questionnaire, and the CAPS Four Week Activity Recall overestimated ($P < 0.05$) daily EE by 62%, 16%, and 11%, respectively, whereas the CAPS Typical Week Activity Recall underestimated ($P < 0.05$) EE by 31%. Both the Harvard Alumni and Five City Project questionnaires overestimated EE in obese and overweight women.

Conclusions: When using 3 of the 4 questionnaire methods, postmenopausal women overestimated EEs. Of all women, obese women overestimated daily EE the most. *Am J Clin Nutr* 2006; 84:230–6.

KEY WORDS Doubly labeled water, physical activity questionnaires, postmenopausal women

INTRODUCTION

In the United States, >65% of adults are now overweight or obese and ≈31% of adults (>61 million people) meet the criteria for obesity (1). Several chronic diseases, such as type 2 diabetes, cardiovascular disease, and certain cancers, are associated with physical inactivity and excessive energy consumption (1). Thus, energy balance and, in particular, its 2 major modifiable components—energy intake and physical activity—have become major public health concerns in the development of chronic diseases.

Assessment of physical activity energy expenditure (EE) is possible through a variety of methods that were developed for different purposes, ranging from physical activity questionnaires (PAQs) to motion sensors and to the measured excretion of isotopic doubly labeled water (DLW) (2). The National Institutes of Health's strategic plan for obesity research identified physical activity EE as methodologically challenging (1). The DLW stable isotope technique is considered the "gold standard" for measuring total daily EE in free-living people (3) because of its accuracy, but it is expensive and impractical for large epidemiologic studies. Self-reported habitual physical activity estimates obtained with PAQs are more practical for use in such epidemiologic studies, but it is important that they be validated against a criterion method such as DLW. In addition to understanding errors associated with the use of self-reported PAQs in population-based studies, this type of validation also has applications in research aimed at optimizing energy balance to reduce disease burden related to overweight and obesity. Few epidemiologic methods for estimating EE with PAQs have been cross-validated against the DLW technique (4–7).

We had the opportunity to evaluate 4 PAQs used in epidemiologic studies against the gold standard DLW method in free-living postmenopausal women. Postmenopausal women constitute a large segment of American women (estimated at ≈40 million in 2000) (8), but the accuracy of EEs obtained with PAQs

[1] From the Department of Epidemiology, The University of Texas MD Anderson Cancer Center, Houston, TX (SM); the US Department of Agriculture, Agricultural Research Service, Beltsville, MD (DJB, BAC, and WSC); the Information Management Services Inc, Silver Spring, MD (CG); the Genetic Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Institutes of Health, Bethesda, MD (PRT); and the Department of Nutrition, Pennsylvania State University, University Park, PA (TJH).

[2] Supported in part by the Intramural Research Program of the National Institutes of Health and the National Cancer Institute.

[3] Reprints not available. Address correspondence to S Mahabir, Department of Epidemiology, CPB4.3247, UT MD Anderson Cancer Center–Unit 1340, 1155 Pressler Boulevard, Houston TX 77030. E-mail: smahabir@mdanderson.org.

Received August 25, 2005.
Accepted for publication March 7, 2006.

compared with those obtained with the DLW method in this population remains unknown.

SUBJECTS AND METHODS

Study design

The present study was part of a randomized, crossover, intervention trial of moderate alcohol supplementation in postmenopausal women ($n = 65$). Details of the study design and procedures have been previously published (9, 10). Briefly, the subjects were assigned to 3 separate 8-wk diet periods during which they consumed a controlled diet and were provided a beverage (orange juice) each day that contained 0, 15, or 30 g alcohol (95% ethanol) in random order. Each subject completed all 3 diet periods, which were separated by 2–5-wk washout periods. All PAQs were administered at baseline.

Subjects

Postmenopausal women were recruited by advertisement from the communities surrounding the Beltsville Human Nutrition Research Center, Beltsville, MD. The eligibility criteria for the women were the following: *1*) aged ≥50 y, *2*) postmenopausal (last menses ≥12 mo before the study started, follicle stimulating hormone >40 000 IU/L, natural menopause, or hysterectomy with ≥1 ovary intact), *3*) not receiving hormone replacement therapy, *4*) not taking prescription medications that may interfere with the study, *5*) willing and able to consume the diet prepared or approved by the Center and no other foods or beverages, *6*) body mass index (BMI) between 90% and 140% of ideal, and *7*) no personal or parental history of alcohol abuse. The subjects were evaluated by a physician and determined to be in good health with no signs or symptoms of any disease or endocrine disorders.

The present study was approved by the National Cancer Institute's Institutional Review Board and the Committee on Human Research of the Johns Hopkins University Bloomberg School of Hygiene and Public Health. All subjects were fully informed of the study requirements and were required to read and sign a consent form detailing the study objectives, risks, and benefits of the study. The subjects were compensated for their participation.

Doubly labeled water method

EE (in kcal/d) was measured with the DLW method. Baseline urine samples from each participant were collected before dosing with oxygen-18 ($H_2{}^{18}O$: 0.14 g/kg body weight) and deuterium (^2H_2O: 0.70 g/kg body weight). Urine was analyzed for deuterium by an automated infrared analysis method in our laboratory and for oxygen-18 by a commercial laboratory subcontractor (Metabolic Solutions Inc, Merrimack, NH). Standards prepared by investigators, but unknown to the commercial laboratory, were used to monitor the performance of the commercial laboratory.

Isotope kinetics were determined by using a multipoint calculation technique (11–13). The deuterium and oxygen-18 zero time intercepts and clearance rates (k_h and k_o) were calculated by using least-squares linear regression on the natural logarithm of the isotope concentration as a function of elapsed time from dose administration. The zero time intercepts were used to determine the isotope pool sizes at the time of the dose. The deuterium and oxygen-18 pool sizes were used to estimate total body water (deuterium pool size/1.04 and oxygen-18 pool size/1.01, respectively). The production rates of carbon dioxide (rCO_2) and water (rH_2O) from the isotope clearance rates (k_h and k_o) and total body water were calculated by the method of Seale et al (11).

Physical activity assessment

The 4 PAQs chosen for the present study have been used in epidemiologic studies, and the essential elements of the questionnaires have been previously published (14–17). These questionnaires were all administered at baseline, and the same interviewer instructed all women in the use of the questionnaires and inspected all completed forms.

PAQ data were analyzed to estimate EE by multiplying duration (in min) by the intensity of each reported activity [in metabolic equivalents (METs)] over the reporting period to measure the average activity level in METs/min. This value was multiplied by the subject's body weight to give an estimation of the total daily EE in kcal. One MET was defined as an oxygen consumption rate of 3.5 mL $\cdot$ kg^{-1} $\cdot$ min^{-1} in adults. Taking the oxygen energy equivalent to be 5 kcal/L (1 L oxygen burns 5 kcal) consumed, this corresponds to 0.0175 kcal $\cdot$ kg^{-1} $\cdot$ min^{-1} (3.5 mL $\cdot$ kg^{-1} $\cdot$ min^{-1} × 0.005 kcal/mL). Comprehensive lists of the energy requirements for specific physical activities are widely available (18, 19). Physical activity EE was computed in kcal/min with the following formula: 0.0175 (kcal $\cdot$ kg^{-1} $\cdot$ min^{-1}/MET^{-1}) × METs × mean weight (in kg) (20). This value was then multiplied by 1440 (the number of minutes in a day) to give kcal/d.

Harvard Alumni questionnaire

This questionnaire(14) was designed to primarily recall leisure-time physical activities associated with cardiovascular disease in college alumni, although it also collects information on all activities usually performed in a typical week. We used the latter portion of the questionnaire in our analysis, as follows. The subjects filled out a table that described the numbers of hours per usual weekday and per usual weekend day that they spent doing activities within each of 4 activity levels (sleeping and light, moderate, and strenuous activity). The METs assigned to these activity levels were 1, 2, 4, and 6, respectively. A weighted average METs/h was computed by multiplying each METS level by time (in h), summing the 4 results, and dividing by the total hours in a week.

Five City Project questionnaire

This questionnaire(15) has 14 items that measures the hours spent sleeping (1 MET) and hours performing moderate (4 METs), hard (6 METs), and very hard (10 METs) intensity physical activities over a week. Weekday and weekend activities are measured separately. Time spent in light intensity activities (2 METs) was computed as follows: 24 h – (time spent sleeping + time in moderate intensity activity + time in hard intensity activity + time in very hard intensity activity). As with the HA questionnaire, a weighted METs/h average was computed and converted to kcal/d.

TABLE 1
Characteristics of the subjects[1]

Characteristic	n (%)	Mean ± SD	Median (interquartile range)	Range
Age (y)	—	59.9 ± 7.5	58 ± 10	49.2–78.8
BMI (kg/m^2)	—	27.7 ± 5.6	27.4 ± 7.6	17.7–42.5
Lean body mass (g)[2]	—	39 512 ± 4895	38 909 ± 6231	29 651–53 548
Percentage body fat (%)[2]	—	41.2 ± 8.6	42.1 ± 11.6	17.8–55.7
Race				
White	49 (75)	—	—	—
Black	12 (18)	—	—	—
Asian	2 (3)	—	—	—
Other	2 (3)	—	—	—

[1] $n = 65$.
[2] Dual-energy X-ray absorptiometry scans were available for only 53 subjects. Body composition was measured by pencil-beam dual-energy X-ray absorptiometry (model DPX-L; Lunar Corp, Madison, WI). The subjects were placed in a supine position with arms and legs close to the body for a whole-body scan, according to the manufacturer's recommended protocol. Whole-body and regional lean mass (mass of bone and nonfat soft tissue) and fat mass were measured by using the manufacturer's algorithm (software version 1.33).

Cross-Cultural Activity Participation Study Four Week Activity Recall and Typical Week Activity Recall questionnaires

The Cross Cultural Activity Participation Study (CAPS) questionnaires were designed to measure the types of moderate and vigorous intensity physical activities performed by minority women aged ≥40 y (17, 21). Two CAPS questionnaires exist: the CAPS Four Week Activity Recall and the CAPS Typical Week Physical Activity Recall. The CAPS Four Week Activity Recall questionnaire identifies specific categories of activities: walking; going places; walking up stairs; climbing ladders; inside housework; outside housework; lawn, garden, farm, or ranch activities; taking care of others; dancing; performing arts, crafts, or musical activities; sport, conditioning, or recreational activities; and working for pay. For each of these categories, the subjects were required to recall how many times the activity was done during the past month, how many times per day the activity was done, and, on average, how many minutes or hours the activity was done each time.

The CAPS Typical Week Activity Recall also lists specific categories of activities: household chores; lawn, yard, garden, or farm work; taking care of others; walking for nonwork reasons; transportation activities; dancing and sports; conditioning activities; leisure activities; occupational activities; volunteer work; infrequent activities done only once or twice in the past 4 wk; important activities done at special times of the year; and other activities. For each of the specified activities, the subjects were required to answer "yes," "no," or "infrequent" and to specify the days of the week that the activity was usually done; if the subjects answered "yes," they were required to recall the minutes or hours per day that the activity was done. As with the other questionnaires, a weighted METs/h average was computed and converted to kcal/d for both CAPS questionnaires.

Statistical analysis

Statistical analyses were performed with SAS for UNIX version 8.0 (SAS Institute, Cary, NC). Means and SDs were calculated to describe the subject characteristics. The estimated EE data from the questionnaires were generally not normally distributed; therefore, medians and interquartile ranges are used to describe these results. Spearman correlations between EE obtained with DLW and those estimated with the PAQs were determined. Simple log transformation normalized the distributions of the questionnaire results; therefore, comparisons between the DLW and the PAQs were accomplished by regressions of the log-transformed measures. Signed-rank tests were used to evaluate the median differences between DLW and the questionnaires.

Effect modification by BMI and other factors (eg, ethnicity or age) was assessed by including the individual factor (eg, BMI) and its cross-product term with the continuous estimate of EE from each of the PAQs (eg, EE questionnaire 1 × BMI) in multivariate models. For these analyses, BMI was coded both as a continuous variable and as a scored (score = 1, 2, or 3) variable and tested both ways in separate models for each questionnaire. Separate multivariate models with EE estimated by DLW as the dependent variable were constructed for each PAQ. The P values associated with the interaction terms and the change in R^2 values between the full models and the models without the interaction term were used to identify potentially important associations. No evidence of either confounding or effect modification by age or ethnicity was observed for any of the questionnaires (data not shown). For BMI, we observed a marginally significant effect modification for the CAPS Typical Week Activity Recall questionnaire and a stronger effect for the CAPS Four Week Activity Recall questionnaire, particularly for the scored BMI variable ($P = 0.04$). Thus, we reported results stratified by BMI and included data for all questionnaires.

We also examined median differences between estimated EE from each questionnaire and from the DLW method within subcategories of age (<60 or ≥60 y), race (white or other), and BMI (normal weight, overweight, or obese). Bland-Altman plots were created to compare the differences between energy intake obtained from the questionnaires and EE obtained with the DLW method. Significance was set at $P = 0.05$, and all P values were two-sided.

RESULTS

The women had a broad range of age, BMI, lean body mass, and percentage body fat (**Table 1**). Of the 65 women who par-

TABLE 2

Spearman correlation coefficients between daily energy expenditure estimated from physical activity questionnaires and from doubly labeled water[1]

Method	HA	FCP	CAPSTW	CAPSFWR	DLW
Harvard Alumni (kcal/d)	1.00				
Five City Project (kcal/d)	0.77[2]	1.00			
CAPS Typical Week Activity Recall (kcal/d)	0.38[2]	0.33[2]	1.00		
CAPS Four Week Activity Recall (kcal/d)	0.32[2]	0.40[2]	0.73[2]	1.00	
Doubly labeled water (kcal/d)	0.36[2]	0.47[2]	0.16	0.15	1.00

[1] HA, Harvard Alumni; FCP, Five City Project; CAPS, Cross-Cultural Activity Participation Study; CAPSTW, CAPS Typical Week Activity Recall; CAPSFWR, CAPS Four Week Activity Recall; DLW, doubly labeled water.

[2] $P < 0.05$.

ticipated in the study, 21 (32%) were normal weight, 26 (40%) were overweight, and 18 (28%) were obese.

Spearman correlations between EEs obtained with the DLW method and those obtained with the PAQs are shown in **Table 2**. The correlations with EE from the DLW method were significant only for the Harvard Alumni and Five City Project questionnaires.

Results of our study show that, compared with the DLW method, the Harvard Alumni, Five City Project, and the CAPS Four Week Activity Recall PAQs all overestimated EE (ie, the percentage differences in the medians were 62%, 16%, and 11%, respectively; **Table 3**). In contrast, the CAPS Typical Week Activity Recall underestimated EE by 31% compared with the DLW method. For all 4 PAQs, these differences in the medians compared with the DLW method were statistically significant (Wilcoxon's signed-rank tests).

We found substantial differences between questionnaires' ability to estimate EE when we examined the women's questionnaire data within BMI categories (ie normal, overweight, and obese women; **Table 4**). The Harvard Alumni questionnaire overestimated EE for all 3 BMI categories, but was best for normal-weight women (47% overestimate), worse for overweight women (65% overestimate), and the worst for obese women (89% overestimate). The Five City Project questionnaire performed substantially better than the Harvard Alumni questionnaire; it was quite precise in normal weight women, only overestimated EE in overweight women by 10%, and overestimated EE in obese women by 32%. The CAPS Four Week Activity Recall also overestimated EE, but only in obese women (by 32%). The underestimation of EE observed for the CAPS Typical Week Activity Recall was evident only in normal-weight (41% underestimate) and overweight (34% underestimate) women. No significant differences between EE estimated by the questionnaires and that obtained with the DLW method were observed in the other subcategories examined (age or race; data not shown).

Graphic comparisons between the DLW and PAQ methods with the use of Bland-Altman plots are shown in **Figures 1–4**. A major positive trend reflecting the substantial overestimation previously noted for the Harvard Alumni questionnaire is shown in the plot in Figure 1. Trends toward overestimation of EE by the Five City Project and the CAPS Typical Week Activity Recall are shown in Figures 2 and 3, respectively. A negative trend indicative of underestimation of EE by the CAPS Four Week Activity Recall is shown in Figure 4.

TABLE 3

Summary of daily energy expenditure (EE) as measured by the doubly labeled water (DLW) method and 4 physical activity questionnaires[1]

Method	Median (interquartile range)	Difference in medians[2]	R^2
	kcal/d		
Basal metabolic rate[3]	1355 (203)	—	—
DLW	2560 (840)	—	—
Harvard Alumni	4149 (1566)	1589[4]	0.11[5]
Five City Project	2970 (1532)	410[4]	0.16[5]
CAPS Typical Week Activity Recall	1760 (1262)	−800[4]	0.02[5]
CAPS Four Week Activity Recall	2851 (3129)	291[4]	0.03[5]

[1] $n = 65$. CAPS, Cross-Cultural Activity Participation Study.

[2] Difference between the median EE from DLW (gold standard) and the median EE from other methods (test). A positive value indicates an overestimation of EE and a negative value indicates an underestimation of EE.

[3] Basal metabolic rate (BMR) was estimated with the following Harris-Benedict equation for women: BMR = 655.1 + [(weight, in kg × 9.563) + (height, in cm × 1.85) − (age, in y × 4.676)].

[4] Significantly different from zero, $P < 0.05$ (Wilcoxon signed-rank test).

[5] Regression of each log-transformed measure against log-transformed DLW.

DISCUSSION

Epidemiologic studies generally require large populations to draw conclusions about associations between disease and exposures; thus, PAQs have become the method of choice in these studies (22–25). Accurate measurement of habitual physical activity is, therefore, important to understand the relations between physical activity and health. Our study is one of only a few studies (4, 6, 7) to simultaneously compare different PAQs against a criterion method (DLW method) to estimate EE in free-living persons, and it is the first to do so in postmenopausal women. Measurements of EE with the DLW method are commonly accepted as the criterion standard for comparing PAQs (4, 5, 26, 27). To the best of our knowledge, except for the Five City Project questionnaire (4, 27), there are no prior reports of DLW validation for the Harvard Alumni or CAPS questionnaires evaluated here.

Compared with our gold standard of the DLW method, our data showed that there were substantial differences in how well these various questionnaires correlated with the DLW method as well as how well they estimated absolute kcal expended per day. The 2 CAPS questionnaires had particularly low correlation coefficients (0.15 and 0.16) compared with the Harvard Alumni

TABLE 4
Summary of daily energy expenditure (EE) by BMI categories, as measured by the doubly labeled water (DLW) method and 4 physical activity questionnaires[1]

	Normal-weight women ($n = 21$)		Overweight women ($n = 25$)		Obese women ($n = 18$)	
Method	Median (interquartile range)	Difference in medians[2]	Median (interquartile range)	Difference in medians[2]	Median (interquartile range)	Difference in medians[2]
	kcal/d		*kcal/d*		*kcal/d*	
Basal metabolic rate	1237 (135)	—	1355 (125)	—	1577 (184)	—
DLW	2357 (807)	—	2665 (631)	—	2730 (1185)	—
Physical activity questionnaire						
Harvard Alumni	3461 (688)	1104[3]	4409 (1015)	1744[3]	5157 (1980)	2427[3]
Five City Project	2324 (339)	−33	2961 (865)	296[3]	4017 (1474)	1287[3]
CAPS Typical Week Activity Recall	1397 (1102)	−960[3]	1747 (1148)	−918[3]	2394 (1738)	−336
CAPS Four Week Activity Recall	2822 (3205)	465	2590 (1474)	−75	3613 (3441)	883[3]

[1] $n = 65$. CAPS, Cross-Cultural Activity Participation Study.
[2] Difference between the median EE obtained with DLW (gold standard) and the median estimated with other methods (test). A positive value indicates an overestimation of EE and a negative value indicates an underestimation of EE.
[3] Significantly different from zero, $P < 0.05$ (Wilcoxon signed-rank test).

(0.36) and Five City Project (0.47) questionnaires. These differences almost certainly reflect that both the Harvard Alumni and Five City Project questionnaires account for all 24 h in a day, including sleep time, whereas the CAPS questionnaires inquire only about light, moderate, and vigorous activity but do not account for sleep time. Although careful assessment of light, moderate, and vigorous activity may be the preferred exposure assessment for some epidemiologic studies that focus on the intensity of physical activity, this approach suffers if the exposure of interest is total EE.

On the absolute scale, the EE estimated with the Harvard Alumni questionnaire differed the most from that obtained with the DLW method (62% overestimation), followed by the CAPS Typical Week Activity Recall (31% underestimation), the Five City Project questionnaire (16% overestimation), and the CAPS Four Week Activity Recall (11% overestimation). One possible reason for the substantial misreporting noted for the Harvard Alumni questionnaire is that this questionnaire was designed to capture only leisure-time and not total physical activity. Arguably, however, this might be expected to result in under- rather than overestimation, as seen here. The underestimation of total EE obtained with the CAPS Typical Week Activity Recall was expected, because this questionnaire only inquires about physical activity for a relatively small part of the day. Overestimation of total EE with the CAPS Four Week Activity Recall, however, was unexpected, and we have no cogent explanation of this result. Although varying in degree, there was misreporting of total EE for all 4 PAQs assessed here. Thus, for epidemiologic studies of free-living postmenopausal women, caution must be exercised if these questionnaires are used in estimating EE because of accuracy issues.

An interesting finding from our study is that heavier women appeared to overestimate EE more than did leaner women in 3 of the 4 questionnaires evaluated. This finding is supported by previous studies that also showed that overweight and obese persons tended to overestimate physical activity EE (28–31).

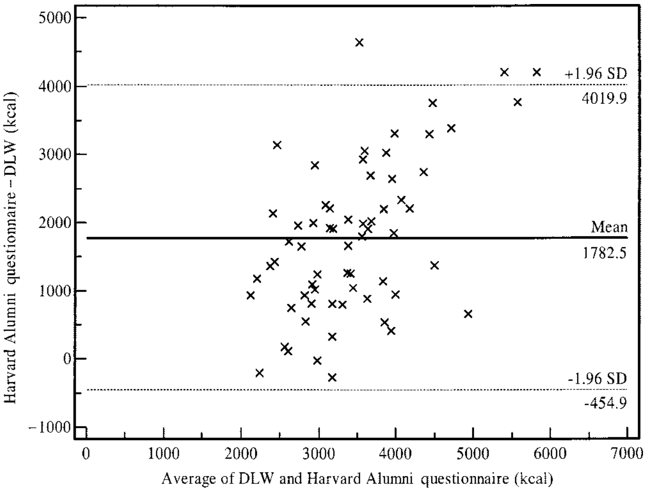

FIGURE 1. Bland-Altman plots of the differences between energy expenditure estimated from the Harvard Alumni physical activity questionnaire and from the doubly labeled water (DLW) method.

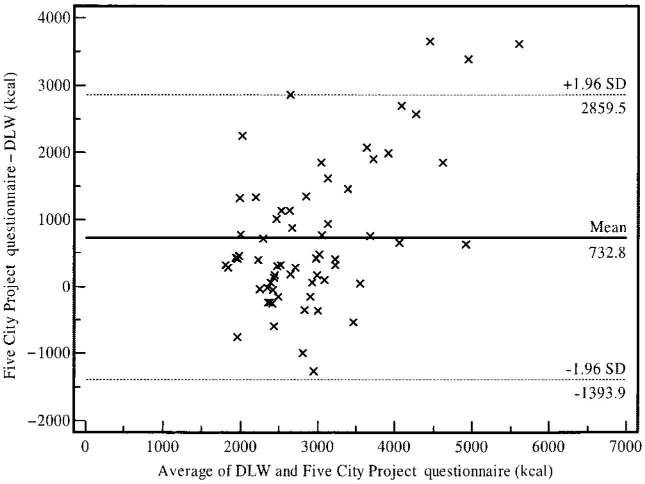

FIGURE 2. Bland-Altman plots of the differences between energy expenditure estimated from the Five City Project physical activity questionnaire and from the doubly labeled water (DLW) method.

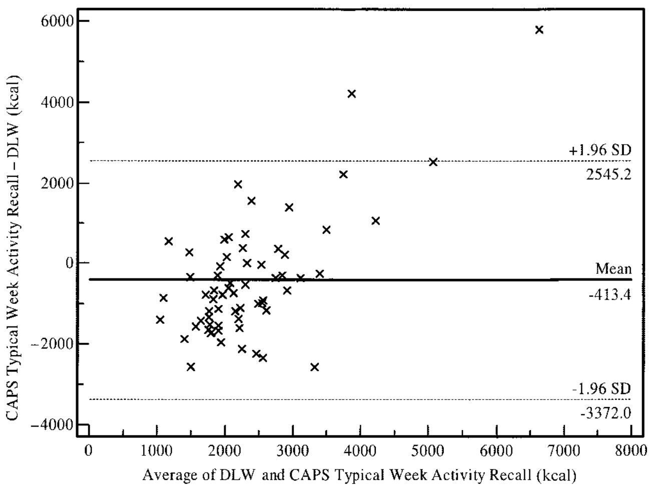

FIGURE 3. Bland-Altman plots of the differences between energy expenditure estimated from the Cross-Cultural Activity Participation Study (CAPS) Typical Week Activity Recall physical activity questionnaire and from the doubly labeled water (DLW) method.

Previous reports that compare the EE estimated from the Five City Project questionnaire to that obtained with the DLW method used different populations than the one we examined here and found different results. Conway et al (4), for example, studied a population of free-living men ($n = 24$) and found that the Five City Project questionnaire significantly ($P < 0.05$) overestimated EE by 31% compared with the DLW method. Washburn et al (27) evaluated a population of young men ($n = 17$) and women ($n = 29$) and found that the estimates of EE from the Five City Project questionnaire and that obtained with the DLW method were not significantly different. In our population of postmenopausal women, we found that the Five City Project questionnaire significantly overestimated EE by 30% compared with the DLW method, making our findings comparable to those of Conway et al (4).

Limitations of the approach we used here merit mentioning. Although the questionnaires assessed habitual physical activity, a single measurement may not be sufficient to accurately capture the range of activities afforded by alternative approaches such as multiple recalls. Self-administration of PAQs, even with prior detailed instruction and a brief interview, may result in bias. Random variation over time in the DLW measurements likely contributed to some of the variation we observed. Also, we did not measure resting EE (ie, basal metabolic rate) or account for the thermic effect of foods, although this is typically relatively small.

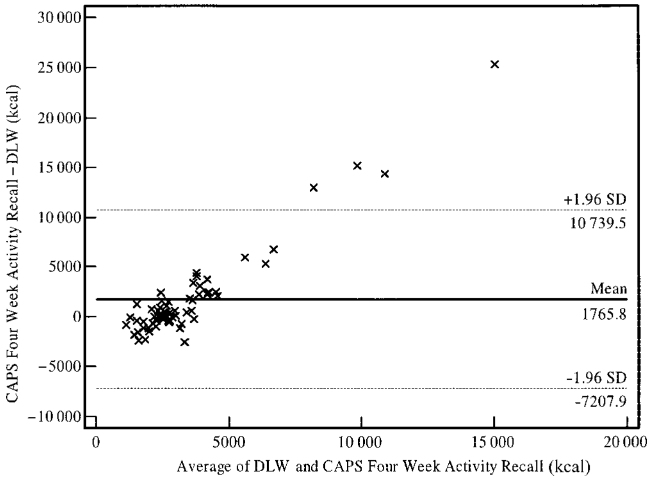

FIGURE 4. Bland-Altman plots of the differences between energy expenditure estimated from the Cross-Cultural Activity Participation Study (CAPS) Four Week Activity Recall physical activity questionnaire and from the doubly labeled water (DLW) method.

In conclusion, we observed substantial misreporting errors in this sample of postmenopausal women when EEs estimated from 4 PAQs were compared with that obtained with the DLW method. In addition, we observed that overweight and obese women tended to overestimate physical activity EE in several of the questionnaires more than did normal-weight women. Taken together, our data suggest that these questionnaires may not be appropriate in assessing total daily EE in free-living postmenopausal women. Because the use of PAQs to accurately estimate EE depends on the ability of the subjects to correctly estimate time spent in different activities of varying intensities, it is clear that additional work aimed at understanding misreporting of habitual physical activities is needed.

We thank Barbara Ainsworth and colleagues for generously providing access to their physical activity questionnaires as well as providing guidance in the administration and programming for scoring of these questionnaires.

SM was involved with the data analysis, data interpretation, and writing of the manuscript. CG was involved in the data analysis. DJB, BAC, PRT, and TJH were involved with the conceptualization, design, execution, and interpretation of the data. WSC was responsible for the management of the study. All the authors contributed to the final manuscript. None of the authors declared any personal or financial conflict of interest.

REFERENCES

1. US Department of Health and Human Services, National Institutes of Health. Strategic plan for NIH obesity research. A report of the NIH Obesity Research Task Force. NIH Publication No. 04-5493. 2004.
2. Keim N, Blanton C, Kretsch M. America's obesity epidemic: measuring physical activity to promote an active lifestyle. J Am Diet Assoc 2004;104:1398–409.
3. Schoeller D, Ravussin E, Schutz Y, Acheson K, Baertschi P, Jequier E. Energy expenditure by doubly labeled water: validation in humans and proposed calculation. Am J Physiol 1986;250:R823–30.
4. Conway J, Seale J, Jacobs D Jr, Irwim M, Ainsworth B. Comparison of energy expenditure estimates from doubly labeled water, a physical activity questionnaire, and physical activity records. Am J Clin Nutr 2002;75:519–25.
5. Conway J, Irwim M, Ainsworth B. Estimating energy expenditure from the Minnesota leisure time physical activity and Tescumseh occupational activity questionnaires—a doubly labeled water validation. J Clin Epidemiol 2002;55:392–9.
6. Racette S, Schoeller D, Kushner R. Comparison of heart rate and physical activity recall with doubly labeled water in obese women. Med Sci Sports Exerc 1995;27:126–33.
7. Starling R, Matthews D, Ades P, Poehlman E. Assessment of physical activity in older individuals: a doubly labeled water study. J Appl Physiol 1999;86:2090–6.
8. Menopause Practice: A Clinician's Guide. Cleveland, OH: The North American Menopause Society, 2004.
9. Baer D, Judd J, Clevidence B, et al. Moderate alcohol consumption lowers risk factors for cardiovascular disease in postmenopausal women fed a controlled diet. Am J Clin Nutr 2002;75:593–9.
10. Dorgan J, Baer D, Albert P, et al. Serum hormones and the alcohol-breast cancer association in postmenopausal women. J Natl Cancer Inst 2001;93:710–5.
11. Seale J, Rumpler W, Conway J, Miles C. Comparison of doubly labeled water, intake balance, and direct- and indirect-calorimetry methods for measuring energy expenditure in adult men. Am J Clin Nutr 1990;52:66–71.

12. Seale J, Miles C, Bodwell C. Sensitivity of methods for calculating energy expenditure by use of doubly labeled water. J Appl Physiol 1989;66:644–53.
13. Seale J, Conway J, Canary J. Seven-day validation of doubly labeled water method using indirect calorimetry. J Appl Physiol 1993;74:402–9.
14. Paffenbarger RJ, Wing A, Hyde R. Physical activity as an index of heart attack in college alumni. Am J Epidemiol 1978;108:161–75.
15. Sallis J, Haskell W, Wood P, et al. Physical activity assessment methodology in the Stanford Five-City Project. Am J Epidemiol 1985;121: 91–106.
16. Taylor H, Jacobs D, Schucker B, Knudsen J, Leon A, DeBacker G. A questionnaire for the measurement of leisure-time physical activities. J Chronic Dis 1978;31:741–55.
17. Ainsworth B, Irwin M, Addy C, Whitt M, Stolarczyk L. Moderate physical activity patterns of minority women: the Cross-Cultural Activity Participation Study. J Womens Health Gend Based Med 1999;8:805–13.
18. Ainsworth B, Haskell W, Leon A, et al. Compendium of physical activities: classification of energy costs of human physical activities. Med Sci Sports Exerc 1993;25:71–80.
19. WIlson P, Paffenbarger R Jr, Morris J, Havlik R. Assessment of methods for physical activity and physical fitness in population studies: a report of NHLBI workshop. Am Heart J 1986;111:1117–93.
20. Powers S, Howley E. Exercise physiology: theory and applications to fitness and performance. 3rd ed. Boston, MA: McGraw Hill, 1993: 481–4.
21. Whitt M, Levin S, Ainsworth B, Dubose K. Evaluation of a two-part survery item to assess moderate physical activity: the Cross-Cultural Activity Participation Study. J Womens Health 2003;12:203–11.
22. Lamonte M, Ainsworth B. Quantifying energy expenditure and physical activity in the context of dose response. Med Sci Sports Exerc 2001;33: S370–8.
23. Albanes D, Conway J, Taylor P, Moe P, Judd J. Validation and comparison of eight physical activity questionnaires. Epidemiol 1990;1:65–71.
24. Shephard R. Limits to the measurement of habitual physical activity by questionnaires. Br J Sports Med 2003;37:197–206.
25. Jacobs D Jr, Ainsworth B, Hartman T, Leon A. A simultaneous evaluation of 10 commonly used physical activity questionnaires. Med Sci Sports Exerc 1993;25:81–91.
26. Hoyt R, Jones T, Stein T, et al. Doubly labeled water measurement of human energy expenditure during strenuous exercise. J Appl Physiol 1991;71:16–22.
27. Washburn R, Jacobsen D, Sonko B, Hill J, Donnelly J. The validity of the Stanford seven day physical activity recall in young adults. Med Sci Sports Exerc 2003;35:1374–80.
28. Jackicic J, Polley B, Wing R. Accuracy of self-reported exercise and the relationship with weight loss in overweight women. Med Sci Sports Exerc 1998;30:634–8.
29. Irwin M, Ainsworth B, Conway J. Estimation of energy expenditure from physical activity measures: determinants of accuracy. Obes Res 2001;9:517–25.
30. Buchowski M, Townsend K, Chen K, Acra S, Sun M. Energy expenditure determined by self-reported physical activity is related to body fatness. Obes Res 1999;7:23–33.
31. Lichtman S, Pisarska K, Berman E, et al. Discrepancy between self-reported and actual caloric intake and exercise in obese subjects. N Engl J Med 1992;327:1893–8.